SCANDINAVIAN JOURNAL OF MEDICINE & SCIENCE IN SPORTS

Physical fitness and academic performance in youth: A systematic review

Journal:	*Scandinavian Journal of Medicine and Science in Sports*
Manuscript ID	SJMSS-R-299-16.R3
Manuscript Type:	Review Article
Date Submitted by the Author:	12-Aug-2016
Complete List of Authors:	Santana, Carla Caroliny; Universidade de Pernambuco, Azevedo, Liane; Teesside University, School of Health and Social Care Cattuzzo, Maria; Universidade de Pernambuco, Physical Education Hill, James; University of Colorado, Anschutz Center for Health and Wellness Andrade, Leylane ; Universidade de Pernambuco, Physical Education Prado, Wagner; Universidade Federal de Sao Paulo, Human Moviment Sciences
Keywords:	Physical education, Physical activity, Physical fitness, School performance, School attendance, Students

Physical fitness and academic performance in youth: A systematic review

Carla Caroliny A. Santana, MS[1], Liane B. Azevedo, PhD[2], Maria T. Cattuzzo PhD[1], James O. Hill, PhD[3], Leylane P. Andrade[1], and Wagner L. Prado, PhD[1,4]

Affiliations: [1]Physical Education Post Graduate Program, University of Pernambuco, Recife, Brazil; [2]School of Health and Social Care, Teesside University, Middlesbrough, England; [3]University of Colorado, Denver, EUA; and [4]Department of Human Movement Sciences, Federal University of São Paulo, Santos, Brazil.

Address correspondence to: Wagner Luiz do Prado, Department of Human Moviment Sciences, Federal University of São Paulo, Rua Silva Jardim, 136, postal code: 11015-020, Santos, São Paulo, Brazil, [wagner.prado@unifesp.br], phone: +551338783874

Short title: Physical fitness and academic performance

Abstract

Physical fitness (PF) is a construct of health and skill related attributes which have been associated with academic performance (AP) in youth. This study aimed to review the scientific evidence on the association among components of PF and AP in children and adolescents. A systematic review of articles using databases PubMed/Medline, ERIC, LILACS, SciELO and Web of Science was undertaken. Cross-sectional and longitudinal studies examining the association between at least one component of PF and AP in children and adolescents, published between 1990 and June 2016, were included. Independent extraction of articles by 2 authors using predefined data fields was done. From a total of 45 studies included, 25 report a positive association between components of PF with AP and 20 describe a single association between cardiorespiratory fitness (CRF) and AP. According to the STROBE guidelines: 12 were classified as low, 32 as medium risk and 1 as high risk of bias. Thirty-one studies reported a positive association between AP and CRF, six studies with muscular strength, three studies with flexibility and seven studies reported a positive association between clustered of PF components and AP. The magnitude of the associations is weak to moderate (β = 0.10 to 0.42 and odds = 1.01 to 4.14). There is strong evidence for a positive association between CRF and cluster of PF with AP in cross-sectional studies; and evidence from longitudinal studies for a positive association between cluster of PF and AP; the relationship between muscular strength and flexibility with AP is uncertain.

Keywords: Physical education, physical activity, physical fitness, school performance, school attendance, students.

Physical fitness (PF) is operationalized as a set of measurable health and skill related attributes including cardiorespiratory fitness (CRF), muscular strength and endurance, body composition, flexibility, balance, agility, coordination, reaction time and power (Caspersen et al., 1985; American College of Sports Medicine, 2010).

Besides the well-described association between PF with a cluster of metabolic risk factors in children, adolescents and adults (Ruiz et al., 2006; Ortega et al., 2008; Kvaavik et al., 2009; Zagout et al., 2016), there is increasing evidence of an association between PF and academic performance (AP) (Grissom, 2005; Chomitz et al., 2009). This could be due to PF's positive effects on cognitive function and performance in attention tasks, or by its effects on depression, stress and sleep quality (Blumenthal et al., 2009; Chang & Chen, 2015; Latorre et al., 2015). Cognition and motor skills in children have several common underlying processes, such as sequencing, monitoring, and planning which may positively influence AP (Roebers & Kauer, 2009; Van der Fels et al., 2015).

AP refers to a child´s success and performance in school and can be measured by grade point averages, as a cluster of achievement tests or, using specific tests for reading or arithmetic skills, such as reading speed, fluency and comprehension, and the ability to solve logical/arithmetic problems (Donnelly et al., 2016).

The components of PF with documented potential for improving health are CRF, muscular strength, and motor ability, each of which may have different effects on the brain (Ruiz et al., 2009; Ruiz et al., 2011). CRF is supposed to induces angiogenesis in the motor cortex and increases blood flow, improving brain vascularization which could affect cognitive performance (Hillman et al., 2008). Muscular strength may alter spinal motoneuron excitability and induce synaptogenesis within spinal cord (Adkins et al., 2006). It is probable that these adaptations in brain may be mediated by physical activity (Hillman et al., 2008). Some studies have examined the association of CRF and muscular strength with AP separately (Aberg et al., 2009; Roberts et al., 2010; Chen et al., 2013; Coe et al., 2013) or clustered (Grissom, 2005; Castelli et al., 2007; Chomitz et al., 2009). However, the results are conflicting and have not been systematically reviewed until now. It has been postulated that body composition is associated with AP (Huizinga et al., 2008; Schwimmer et al., 2003). It is not clear if AP can be directly mediated by fatness or if it is influenced by factors associated with excessive fatness, such as poor self-esteem (Fan et al., 2010), anxiety/depression (Goldfield et al., 2010), teasing and social rejection (Gunnarsdottir et al., 2011), poor attendance in school classes (Pan et al., 2013) and poor PF (Datar et al., 2004; Shore et al., 2008).

Keeley & Fox (2009) reviewed the impact of physical activity (PA) and fitness on academic achievement and cognitive performance in children and found a weak relationship between PF (mainly CRF) and academic achievement. It is important to note that only seven studies were reviewed and a qualitative analysis of the studies was not performed. Recently, Haapala (2013) in a narrative review described a positive association between CRF and AP in children. However, this review was not systematic and did not considered all components of PF. Therefore, the aim of this systematic review was to examine the scientific evidence on the associations between PF´s components and AP in children and adolescents.

Materials and methods

This systematic review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria (Moher et al., 2009).

Criteria for Considering Studies for this Review

Cross-sectional and longitudinal studies were considered for this review. Eligible studies were: (1) published in English; (2) conducted among children and adolescents aged <20 years; (3) those that reported an association between CRF, muscle strength or endurance (assessed either field tests or in a laboratory), flexibility and AP (assessed by objective measures, such as, grade point average, standardized test scores and teacher rated achievement); (4) published prior to June 2016. Review articles, meta-analysis, validation studies, conference abstracts, monographs, dissertations, theses, commentaries, brief reports and studies conducted with special populations, such as children or adolescents with developmental disabilities, developmental delays or cognitive impairment were not included.

Search Methods for Identification of Studies

A two-phase search strategy was used to identify potential studies. First, five electronic data bases were systematically searched: PubMed/Medline, ERIC, LILACS, SciELO and Web of Science. The search for descriptors and terms was performed with the MeSH – Medical Subject Headings, through the portal of the U.S. National Library of Medicine (NLM) and DeCS – Descriptors in Health Science from the database Virtual Health Library (Biblioteca Virtual da Saúde- BVS). Search strategies included the combination of variations between two groups of key-words/terms including, but not limited to the following examples: (1) PF (“physical fitness”, “physical endurance”, “cardiorespiratory fitness”, “physical conditioning”, “muscular strength”, “muscular endurance”, “muscular resistance”, flexibility, pliability); and (2) AP (“academic achievement”, “academic performance”, “school performance”, “educational status” "attendance school"). Terms were combined using the logical operators available as search tools. A detailed description of the Pubmed/Medline search is shown in Appendix 1. The search strategy was adapted for the other databases.

In the second phase, the titles and abstracts for potentially relevant articles were screened by two researchers (CCA and LPA). Afterwards, a full copy of the papers was obtained for those that met the initial screening criteria and then were fully examined by two researchers. If the two reviewers were unable to reach a consensus a third researcher (WLP) was consulted. Furthermore, we examined references cited by each of these articles in an attempt to identify other potential studies for inclusion.

Data extraction

The following data were extracted from all eligible articles: (a) year of publication, (b) study design, (c) location, (d) sample size, (e) age, (f) school year (grades), (g) main outcomes (PF components and AP), (h) the main finding (i), and control variables including socioeconomic status (SES), educational level of parents, physical activity level, health status and others.

Risk of bias

The risk of bias within studies was assessed using the adapted Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement according to Lubans et al. (2010). A score of 0 (absent or inadequately described) or 1 (present and explicitly described) was assigned to each of the following questions: (1) Does the study describe the eligibility criteria for participant selection? (2) Are participants randomly chosen from the population? (3) Does the study report the sources and details of the method for measuring academic performance and did the instrument have acceptable reliability? (4) Does the study report the sources and details of the assessment methods for measuring obesity and did the instrument have acceptable reliability? (5) Does the study report the power calculation and use a statistical method that is adequate to test the hypothesis? (6) Does the study report the number of participants for each outcome measurement and does this number represent at least 80% of the total sample?

A score for each article ranged from zero to six points. Studies with ≤ 2 were considered high risk of bias, studies that achieved 3-4 points were classified as medium risk and those that had scores of 5- 6 were classified as low risk of bias. Two independent researchers performed the quality assessment (CCAS and LPA), and disagreements were solved by the lead researcher (WLP).

The judgment of overall scientific evidence was based on Lubans et al. (2010) using the percentage of studies that reported statistically significant relationship but also taking into account the risk of bias: (a) *No Association,* if less than 33% of the studies indicate a significant association between variables; (b) *Uncertain association,* if 34-59% of the studies indicate a significant association between variables; (c) *Positive (or negative) association,* if 60-100% of the studies indicated a significant association (in the same direction); (d) *Strong evidence of positive (or negative) association,* if 60-100% of the studies indicated a significant association between variables (in the same direction) and more than 59% of the studies deemed low risk of bias (score ≥5) found a significant association.

Results

Study selection

The initial search identified a total of 1,404 potential papers. After adjusting for duplicates, 1,369 remained. Of these, 1,315 studies did not meet the criteria after examination of the title and abstract. The full texts of the 54 remaining papers were examined. Reference tracking identified two additional studies (Kantomaa et al., 2013; Hansen et al., 2014), resulting in total of 56 papers. Of these, 11 studies did not meet the inclusion criteria. Two papers did not use an objective measure for AP (Padilla-Moledo et al., 2012; Raine et al., 2013). Two papers were conducted in adults (Saavedra et al., 2008; Valkeinen et al., 2013). One paper was a meta-analysis (Fedewa & Ahn, 2011). Two papers were commentaries (Feiden, 2011; Hupert, 2015). Three were dissertations (Hanna, 2009; Hannigan, 2010; Mobilia-Jones, 2010). One paper was published in Spanish (Cumillaf et al., 2015). Therefore, 45 studies met all inclusion criteria and were included in this systematic review (Figure 1).

Figure 1 here

Characteristics of included studies

Table 1 shows that 77.8% of the studies had a cross-sectional design, 22.2% had prospective design and 51.1% of the studies were conducted in the United States.

Table 1 here

Risk of bias assessment indicated that a single study (2.2%) was classified as high risk (Du Toit et al., 2011), 32 (71.1%) as medium risk (Kim et al., 2003; Castelli et al., 2007; Chomitz et al., 2009; Eveland-Sayers et al., 2009; Wittberg et al., 2009; Roberts et al., 2010; Wittberg et al., 2010; Blom et al., 2011; Davis and Cooper, 2011; London and Castrechini, 2011; Van Dusen et al., 2011; Wingfield et al., 2011; Wittberg et al., 2012; Bass et al., 2013; Chen et al., 2013; Coe et al., 2013; Kantomaa et al., 2013;

Lambourne et al., 2013; Liao et al., 2013; Rauner et al., 2013; Greeff et al., 2014; Haapala et al., 2014; Hansen et al., 2014; Scudder et al., 2014; Srikanth et al., 2014; Torrijos-Niño et al., 2014; Chaddock-Heyman et al., 2015; Desai et al., 2015; Pellicer-Chenoll et al., 2015; García-Hermoso, 2016; Morita et al., 2016; Pindus et al., 2016) and 12 (26.7%) as low risk (Dwyer et al., 2001; Kwak et al., 2009; Welk et al., 2010; Telford et al., 2012; Bezold et al., 2014; Esteban-Cornejo et al., 2014; Janak et al., 2014; Sardinha et al., 2014; Aguilar et al., 2015; Huang et al., 2015; Kalantari et al., 2016; Sardinha et al., 2016) (Table 2).

Table 2 here

Reporting how participants were randomly chosen from the population was the main limiting factor in terms of risk of bias, a criterion that only 8 (17.8%) studies fulfilled (Dwyer et al., 2001; Kim et al., 2003; Kwak et al., 2009; Lambourne et al., 2013; Hansen et al., 2014; Sardinha et al., 2014; Kalantari et al., 2016; Sardinha et al., 2016). This was followed by reporting statistical power as only 13 (28.9%) studies fulfilled this criterion (Dwyer et al., 2001; Kwak et al., 2009; Welk et al., 2010; Wittberg et al., 2010; Blom et al., 2011; Telford et al., 2012; Rauner et al., 2013; Bezold et al., 2014; Esteban-Cornejo et al., 2014; Janak et al., 2014; Aguilar et al., 2015; Huang et al., 2015, Morita et al., 2016). All studies present acceptable standards of quality of reliable measurement from both exposure and outcome variables. However, no study fulfilled all criteria to achieve the maximum quality score.

Participants

The sample size ranged from 42 to 2,550,114 students. Participants were enrolled in 1st grade to senior high school and aged from 6 to18 years old. Seven studies did not report the age of participants (Blom et al., 2011; Wittberg et al., 2012; Coe et al., 2013; Rauner et al., 2013; Bezold et al., 2014; Janak et al., 2014; Pellicer-Chenoll et al., 2015). Studies were conducted in both genders, except Kalantari et al. (2016) which was conducted in boys.

Dependent variable outcome: academic performance

AP was measured by standardized tests to assess knowledge in reading, writing, mathematics and general knowledge in 37 studies (82.2%). Seven studies (16.7%) used measures of AP through by scores given by the teachers (Kim et al., 2003; Kwak et al., 2009; Chen et al., 2013; Kantomaa et al., 2013; Kalantari et al., 2016; Morita et al., 2016; Sardinha et al., 2016). One study (2.4%) only utilized a math test (Huang et al., 2015).

Independent variable outcomes

Ten studies (22.2%) used the FITNESSGRAM test to assess body composition, CRF, muscular strength, muscular endurance and flexibility, (Castelli et al., 2007; Chomitz et al., 2009; Wittberg et al., 2009; Blom et al., 2011; Du Toit et al., 2011; London & Castrechini, 2011; Van Dusen et al., 2011; Bass et al., 2013; Coe et al., 2013; Bezold et al., 2014). Three studies (6.7%) utilized the Alpha battery tests to asses CRF, agility and muscular fitness (Esteban-Cornejo et al., 2014; Torrijos-Niño et al., 2014; García-Hermoso, 2016). Thirty-seven studies (82.2%) examined the association between CRF and AP, seven (15.6%) looked at the association between muscular strength and AP, ten studies (22.2%) between flexibility and AP and eight studies (17.8%) examined the association between PF components clustered and AP.

Twenty studies (44.4%) assessed only CRF (Kwak et al., 2009; Roberts et al., 2010; Welk et al., 2010; Wittberg et al., 2010; Davis & Cooper, 2011; Telford et al., 2012; Wittberg et al., 2012; Kantomaa et al., 2013; Lambourne et al., 2013; Rauner et al., 2013; Haapala et al., 2014; Hansen et al., 2014; Janak et al., 2014; Sardinha et al., 2014; Scudder et al., 2014; Srikanth et al., 2014; Chaddock-Heyman et al., 2015; Huang et al., 2015; Pindus et al., 2016; Sardinha et al., 2016).

Some studies also measured motor skills. Three studies (6.7%) assessed motor proficiency (Du Toit et al., 2011), static balance or manual dexterity (Haapala et al., 2014) and motor ability (Esteban-Cornejo et al., 2014) and 1 study added measures of gross motor skills (Kantomaa et al., 2013).

Main results

According to established criteria of overall scientific evidence (Lubans et al. 2010), 27/30 cross-sectional studies (90%), and 4/7 longitudinal studies (57%) report a positive association between CRF and AP (Table 3). The results suggest strong evidence of positive associations between CRF and AP for cross-sectional studies and uncertain evidence for longitudinal ones.

Sixteen studies (35.5%) investigated the association between muscular strength and AP, 5/14 cross-sectional studies (35.7%) and 1/2 longitudinal studies (50%) reported a positive association between muscular strength with AP. There is uncertain evidence about the association between muscular strength and AP for both cross-sectional and longitudinal studies. Ten studies focused on the association between flexibility and AP (22.2%), 2/8 cross-sectional studies (25%) and 1/2 longitudinal (50%) described a positive association between flexibility and AP. There is no evidence from cross-sectional, and uncertain evidence for longitudinal association between flexibility and AP. Eight studies (17.8%) investigated the association between a clustered of PF components and AP, from this, 4/5 cross-sectional studies (80%) and

3/3 longitudinal studies (100%) reported a positive association. For cross-sectional studies there is a positive association between cluster of PF and AP, for longitudinal studies there is a strong evidence of positive association (table 3).

Table 3 here

Three studies found that the association between CRF and AP is stronger in girls than boys. Eveland-Sayers et al. (2009) verified that time in a 1-mile run test was negatively associated with performance in reading/language (r=-0.31) and math (r=-0.36) in girls. Du Toit et al. (2011) found that relationship between PF components and AP was stronger among girls (CRF r=0.63; knee push-ups r=0.62; wall sitting r=0.54; strength total r=0.62) than boys (CRF r=0.40; knee push-ups r=0.45; wall sitting r=0.16; strength total r=0.51). Van Dusen et al. (2011) showed a higher effect size in the association between CRF and reading skills in girls (effect size = 0.27 -CI:0.25-0.29) compared to boys (effect size = 0.17 - CI:0.15-0.19). A single study described an association between anaerobic power and flexibility with AP only in girls (Liao et al., 2013). Four studies found no or weak association between PF and CRF with AP after adjusting for SES, BMI, motor performance, and screen time (Wingfield et al., 2011; Haapala et al., 2014; Janak et al., 2014; Aguilar et al., 2015). Kantomaa et al. (2013) did not find a relationshihp between compromised motor function in childhood, CRF and AP in adolescents.

The association between PF with AP was assessed by multivariate linear regression analyses in 21 studies (46.6%) (Dywer et al., 2001; Kim et al., 2003; Kwak et al., 2010; Castelli et al., 2007; Roberts et al., 2010; London and Castrechini, 2011; Wingfield et al., 2011; Chen et al., 2013; Coe et al., 2013; Liao et al., 2013; Esteban-Cornejo et al., 2014; Greeff et al., 2014; Haapala et al., 2014; Janak et al., 2014;

Srikanth et al., 2014; Aguilar et al., 2015; Huang et al., 2015; García-Hermoso, 2016; Kalantari et al., 2016; Morita et al., 2016; Pindus et al., 2016), 8 (17.8%) used logistic binomial regression (Chomitz et al., 2009; Welk et al., 2010; Blom et al., 2011; Bass et al., 2013; Sardinha et al., 2014; Torrijos-Niño et al., 2014; Desai et al., 2015; Sardinha et al., 2016). Five (11.1%) studies utilized mixed models regression (Van Dusen et al., 2011; Telford et al., 2012; Lambourne et al., 2013; Rauner et al., 2013; Bezold et al., 2014). One study (2.4%) utilized structural equation to test mediation (Kantomaa et al., 2013), and 1 study (2.2%) used multilevel regression (Hansen et al., 2014). Three studies (6.6%) utilzed analyses of variance (ANOVA) (Wittberg et al., 2009; Wittberg et al., 2012; Chaddock-Heyman et al., 2015). Four studies (8.9%) did simple correlation (Eveland-Sayers et al., 2009; Wittberg et al., 2010; Davis and Cooper 2011; Du toit et al., 2011). One (2.2%) study utilized Student t-test (Scudder et al., 2014), and 1 study (2.2%) utilized the SOM equation (Peliccer-Chenoll et al., 2015).

Few studies provided details of the statistical analysis, when using linear regression analysis and logistic regression (Dwyer et al., 2001; Castelli et al., 2007; Chomitz et al., 2009; Kwak et al., 2010; London & Castrechini, 2011; Chen et al., 2013; Liao et al., 2013; Esteban-Cornejo et al., 2014; Greeff et al., 2014; Sardinha et al., 2014; Morita et al., 2016; Pindus et al., 2016). Another important consideration about the analysis was how the covariates were incorporated in the final model. Considering that the covariates are potentially related to one another, only the study of Esteban-Cornejo et al. (2014) performed the final model analysis using multicollinearity using variance inflation factors (VIFs). Two studies reported analysis of normal distribution of error terms and outliers; however, the authors do not report details of the analysis, such as, influence analysis of the fitted model, studentized residuals, Cook's Distance and distributions of unweighted and weighted models (London & Castrechini, 2011; Pindus et al., 2016).

Most population-based studies showed an association between PF and AP, and these findings can be extrapolated to other populations due external validity. However, the magnitude and power of the results were small in some studies (Dwyer et al., 2001; Roberts et al., 2010; Welk et al., 2010; Blom et al., 2011; Van Dusen et al., 2011; Liao et al., 2013; Rauner et al., 2013; Bezold et al., 2014). Van Dusen et al. (2011) with a sample size of 254,743 students found an association between CRF and performance in reading (effect size = 0.27) and performance in math (effect size = 0.33) in girls, for boys the reported effect size was 0.17 for reading and for math 0.34, respectively. Liao et al. (2013) with size sample of 149, 240 students verified that a 1 Standard Deviation (SD) increase on the anaerobic power and flexibility Z-scores from high school in girls was associated with an increase in the university entrance exam score by 0.018 and 0.010 SD, respectively. Dwyer et al. (2001) found a very weak correlation between time to complete the 50-meter run (r=-0.15), sit-ups repetitions (r=0.14), and distance leaped in the standing long jump (r=0.10) and AP. Welk et al. (2010) with a sample size of 36,835 students also found a weak association between CRF and AP (OR=1.014; CI=1.011-1.016), indicating that for each 1% increase in CRF, AP increases.

Regarding studies with lower size sample, some utilized direct measure to asses CRF and added measures of cognition, electroencephalographic and magnetic resonance imaging, which increases the internal validity of the studies (Davis & Cooper, 2011; Haapala et al., 2014; Scudder et al., 2014; Chaddock-Heyman et al., 2015; Pindus et al., 2016). Although some studies did not use more elaborate statistical tests due to limited sample size, the authors found significant associations between CRF and performance in reading (r=0.29) and math (r=0.25) (Davis & Cooper, 2011). Scudder et al. (2014)

verified that more fit children had greater reading achievement scores (123.1 ± 2.8) compared to less fit children (112.9 ± 2.1), with a larger effect size d= 0.8. A similar finding was observed for spelling (higher fit: 117.6 ± 2.6; lower fit: 108.0 ± 3.0; d= 0.7). Chaddock-Heyman et al. (2015) reported that higher fit children showed decreased cortical thickness in the superior frontal cortex (3.76± 0.15; d=0.62), superior temporal cortex (3.17± 0.24; d= 0.64) and lateral occipital cortex (2.46± 0.10; d= 0.68) when compared to lower fit children (3.85± 0.14; 3.11± 0.18; 2.56± 0.19), respectively. Therefore, these methodological discrepancies should be considered when interpreting the results.

Discussion

From this systematic review, we conclude that there was strong evidence for a positive association between CRF and AP in cross-sectional studies; a positive association with the cluster of PF and AP; and a strong evidence from longitudinal studies for a positive association between cluster of PF and AP. These results support the conclusion of a recent review from Donnelly et al. (2016) demonstrating cross-sectional and longitudinal associations between PF, AP and cognition in youth. However, those authors state that these conclusions should be cautiously interpreted as they are based on cross-sectional, acute/short-term, nonrandomized trials and from randomized trials with a high risk of bias. Many of the studies had relatively small samples or used correlational methodologies that cannot provide evidence of cause and effect. Because most of the existing studies are cross-sectional or correlational, there is a need for intervention studies to determine whether this relationship is causal. If the relationship is causal, there would be a strong reason to promote physical activity in schools as a means of increasing academic achievements.

The quality of existing studies was uneven and future studies could be strengthened by longitudinal research and follow-up assessments for RCT designs to provide a better understanding of the longevity of PF effects on cognition and academic achievement. Although the best evidence will come from RCT designs, in cases where cross-sectional data are still collected, it is recommended that researchers study the contribution of each component of PF on AP in order to determine whether another component of PF is really more important than other.

Most of the studies showed weak to moderate associations between PF and AP; the magnitude ranged from 0.10 (Kim et al., 2003) to 0.42 (Castelli et al., 2007) and odds ratio from 1.01 (Welk et al., 2010) to 4.14 (Blom et al., 2011). The magnitude of the association might not be clinically meaningful at individual level, but might be at population level. This systematic review expands a previous review by Keeley & Fox by including studies since 2009, and by including a qualitative assessment of the studies. We found 45 studies that met our criteria for inclusion in the systematic review. Of these, 30 (66.6%) with medium and low risk reported a positive association between CRF with AP and 7 studies that measured PF as cluster found a positive association between PF and AP.

PF has been associated with improvement in cognitive function in children and better performance in attention tasks. It is possible that those children who have higher cognitive ability may have both better AP and better PF compared to those with lower level of cognitive ability (Hillman et al., 2005; Hillman et al., 2009). This may be related to parents' cognitive ability and SES (Tucker-Drob & Harden, 2012; Heberle & Carter, 2015), which are not taken into account in some studies (Eveland-Sayers et al., 2009; Wittberg et al., 2010; Du Toit et al., 2011; Wittberg et al., 2012; Pellicer-Chenoll et al., 2015).

The mechanism of the association between academic performance and CRF remains unclear. It is hypothesized that improvements in CRF, as induced by changes in physical activity levels, may have a positive influence on cognition mediated by increased levels of brain-derived neurotrophic factor (BDNF) (Hillman et al., 2008). In a systematic review, Knaepen et al. (2010) suggested that the BDNF response to exercise is most probably a phenomenon of what happens centrally, and exercising regularly could induce central effects without elevating peripheral basal BDNF concentration. However, the association between CRF and AP might be influenced by a number of confounding factors affecting CRF. Some studies reported an inverse association between obesity and CRF (Pahkala et al., 2013; Díez-Fernández et al., 2014), and others suggested that due to growth and maturation the relationship between these two variables might be more complex in children and adolescents (Rowland, 2013). It has been reported that CRF and body fat contribute equally to endurance performance in field tests (Rowland et al., 1999).

The relationship between CRF and AP has been investigated mainly by CRF field-based exercise tests (Roberts et al., 2010; Welk et al., 2010; Wittberg et al., 2010; Lambourne et al., 2013; Rauner et al., 2013; Telford et al., 2012; Janak et al., 2014; Sardinha et al., 2014; Srikanth et al., 2014; Huang et al., 2015; Sardinha et al., 2016). There is evidence that the association between CRF and AP might be protocol/test dependent (Dwyer et al., 2001), however it's suggested that this association may be influenced by motor skills and running efficiency (Haapala et al, 2014; Van Der Niet et al., 2014) rather than CRF itself (Kantomaa et al., 2011; Kantomaa et al., 2013).

The different association between boys and girls were also explored in a few studies. Kwak et al. (2009) found that fitness was associated with AP in boys but not in girls. On the other hand, others reported strong associations in girls (Eveland-Sayers et al., 2009; Du Toit et al., 2011; Van Dusen et al., 2011; Liao et al., 2013). Haapala et al. (2014) reported that in girls maturation may, at least partly, explain the relationships of shuttle run performance and AP. Differences could be attributed to the age of the subjects, such as, adolescents (Kwak et al., 2009; Van Dusen et al., 2011; Liao et al., 2013), children (Eveland-Sayers et al., 2009; Du Toit et al., 2011; Haapala et al., 2014), or assessment of PF and sex (Kalantari et eal., 2016).

Considering the results presented in this review, studies are needed to develop effective strategies for improving AP in school age children. Nevertheless, education professionals are under pressure to improve the AP of students and historically, the physical education classes are supposed to be less valuable than other disciplines in an education setting. As a result, many schools have reduced or negleted children's opportunities for PA at school (Chaddock et al., 2010). Reductions in time devoted to PA and an increased amount of time in sedentary behaviors may not only affect weight status but could also reduce HRPF, which could negatively affect AP (Tokmakidis et al., 2006; Brunet et al., 2007; Korsten-Reck et al., 2007; Cantell et al., 2008; Nevill et al., 2009).

The present study has some limitations that should be considered. A literature search was performed only in journals indexed in the electronic databases ERIC, PubMed / Medline, LILACS, Web of Science and SciELO. Therefore, it is possible that some studies were not retrieved. Besides, the inherent limitations of cross-sectional study designs, which were noted the most in this review, preclude making definitive conclusions regarding causality relating to the development of AP and PF. Another important limitation to consider is publication bias (e.g. studies which found no association may not been published), which might affect the findings of this review.

Finally, there was only a small proportion of the studies (28.9%) meeting the criterion of estimating statistical power.

In summary, the findings encourage carefully controlled intervention studies to determine if increasing PF could improve AP of students.

Perspectives

The present review aimed to examine the scientific evidence on the association between PF and AP in children and adolescents. Based on the results of this review, we suggest the focus of future interventions should be directed at promoting PF in an environment that is developmentally appropriate for children and adolescents as this may be the most advantageous path to promote overall functional capabilities, as well as improvements on motor skill due to gender-specific games, equipment and spaces (Hardy et al., 2012). Both PF and motor skill may promote positive and sustainable trajectories of health, cognition and AP leading to long term positive health outcomes.

References

Adkins D, Boychuk J, Remple M, Kleim J. Motor training induces experience-specific patterns of plasticity across motor cortex and spinal cord. J Appl Physiol 2006: 101: 1776-1782.

Aberg MA, Pedersen NL, Torén K, Svartengren M, Bäckstrand B, Johnsson T, Cooper-Kuhn CM, David Aberg N, Nilsson M, Kuhn HG. Cardiovascular fitness is associated with cognition in young adulthood. Proc Natl Acad Sci USA 2009: 106: 20906-20911.

Aguilar MM, Vergara FA, Velásquez EJ, Marina R, García-Hermoso A. Screen time impairs the relationship between physical fitness and academic attainment in children. J Pediatr 2015: 91: 339-345.

American College of Sports Medicine. ACSM's guidelines for exercise testing and prescription. 8th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2010.

Bass RW, Brown DD, Laurson KR, Coleman MM. Physical fitness and academic performance in middle school students. Acta Paediatr 2013: 102: 832-837.

Bezold CP, Konty KJ, Day SE, Berger M, Harr L, Larkin M, Napier MD, Nonas C, Saha S, Harris TG, Stark JH. The effects of changes in physical fitness on academic performance among New York city youth. J Adolesc Health 2014: 55: 774-781.

Blom LC, Alvarez J, Zhang L, Kolbo J. Associations between health-related physical fitness, academic achievement and selected academic behaviors of elementary and middle school students in the state of Mississippi. Journal of Research 2011: 6: 13-19.

Blumenthal JA, Babyak MA, Moore KA, Craighead WE, Herman S, Khatri P, Waugh R, Napolitano MA, Forman LM, Appelbaum M, Doraiswamy PM, Krishnan KR. Effects of exercise training on older patients with major depression. Arch Intern Med 2009: 159: 2349-2356.

Brunet M, Chaput JP, Tremblay A. The association between low physical fitness and high body mass index or waist circumference is increasing with age children: The ''Quebec en Forme'' project. Int J Obes 2007: 31: 637-643.

Cantell M, Crawford SG, Doyle-Baker PK. Physical fitness and health indices in children, adolescents and adults with high or low motor competence. Hum Mov Sci 2008: 27: 344-362.

Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985: 100: 126-131.

Castelli DM, Hillman CH, Buck SM, Erwin HE. Physical fitness and academic achievement in third- and fifth-grade students. J Sport & Exerc Psychol 2007: 29: 239-252.

Chaddock L, Erickson KI, Prakash RS, Kim JS, Voss MW, Vanpatter M, Pontifex MB, Raine LB, Konkel A, Hillman CH, Cohen NJ, Kramer AF. A neuroimaging investigation of the association between aerobic fitness, hippocampal volume, and memory performance in preadolescent children. Brain Res 2010: 1358: 172-183.

Chaddock-Heyman L, Erickson KI, Kienzler C, King M, Pontifex MB, Raine LB, Hillman CH, Kramer AF. The role of aerobic fitness in cortical thickness and mathematics achievement in preadolescent children. Plos One 2015: 10: 1-11.

Chang SP, Chen YH. Relationship between sleep quality, physical fitness and body mass index in college freshmen. J Sports Med Phys Fitness 2015: 55: 1234-1241.

Chen LJ, Fox KR, Ku PW, Taun CY. Fitness change and subsequent academic performance in adolescents. J Sch Health 2013: 83: 631-638.

Chomitz VR, Slining MM, McGowan RJ, Mitchell SE, Dawson GF, Hacker KA. Is there a relationship between physical fitness and academic achievement? Positive results from public school children in the northeastern United States. J Sch Health 2009: 79: 31-37.

Coe DP, Peterson T, Blair C, Schutten MC, Peddie H. Physical fitness, academic achievement, and socioeconomic status in school-aged youth. J Sch Health 2013: 83: 500-507.

Cumillaf AG, Badilla PV, Herrera CF, Mora FC, Herrera BM, Sandoval EM, Muñoz RG, Agüero SD. Asociación entre la condición física, estado nutricional y rendimiento académico en estudiantes de educación física. Nutr Hosp 2015: 32: 1722-1728.

Datar A, Sturm R, Magnabosco JL. Childhood overweight and academic performance: national study of kindergartners and first-grades. Obes Res 2004: 12: 58-68.

Davis CL, Cooper S. Fitness, fatness, cognition, behavior, and academic achievement among overweight children: Do cross-sectional associations correspond to exercise trial outcomes? Prev Med 2011: 52: 65-69.

Desai IK, Kurpad AV, Chomitz VR, Thomas T. Aerobic fitness, micronutrient status, and academic achievement in Indian school-aged children. Plos One 2015: 10: 1-13.

Díez-Fernández A, Sanchez-Lopez M, Mora-Rodriquez R, Notario-Pacheco B, Torrijos Nino C, Martinez-Vizcaino V. Obesity as a mediator of the influence of cardiorespiratory fitness on cardiometabolic risk: a mediation analysis. Diabetes Care 2014: 37: 855-862.

Donnelly JE, Hillman CH, Castelli D, Etnier JL, Lee S, Tomporowski P, Lambourne K, Szabo-Reed AN. Physical activity, fitness, cognitive function, and academic achievement in children: A systematic review. Med Sci Sports Exerc: 2016: 48: 1197-1222.

Du Toit D, Pienaar AE, Truter L. Relationship between physical fitness and academic performance in South African children. S Afr J Res Sport Phys Educ Recreation 2011: 33: 23-35.

Dwyer T, Sallis JF, Blizzard L, Lazarus R, Dean K. Relation of academic performance to physical activity and fitness in children. Pediatr Exerc Sci 2001: 13: 225-237.

Esteban-Cornejo I, Terejo-González CM, Martinez-Gomez D, del-Campo J, González-Galo A, Padilla-Moledo C, Sallis JF, Veiga OL, UP & DOW Study Group. Independent and combined influence of the components of physical fitness on academic performance in youth. J Pediatr 2014: 165: 306-312.

Eveland-Sayers BM, Farley RS, Fuller DK, Morgan DW, Caputo JL. Physical fitness and academic achievement in elementary school children. J Phys Act Health 2009: 6: 99-104.

Fan Y, Li Y, Liu A, Hu X, Ma G, Xu G. Associations between body mass index, weight control concerns and behaviors, and eating disorder symptoms among non-clinical Chinese adolescents. BMC Public Health 2010: 10: 1-12.

Fedewa AL, Ahn S. The effects of physical activity and physical fitness on children's achievement and cognitive outcomes: a meta-analysis. Res Q Exerc Sport 2011: 82: 521-535.

Feiden K. The Texas youth fitness study: looking at school policies as they relate to physical fitness and academic variables. Programs results report. ERIC (Non-Journal). 2011: 1-9.

García-Hermoso. Aerobic capacity as a mediator of the influence of birth weight and school performance. J Dev Orig Health Dis 2016: 7: 337-341.

Goldfield GS, Moore C, Henderson K, Buchholz A, Obeid N, Flament MF Body dissatisfaction, dietary restraint, depression, and weight status in adolescents. J Sch Health 2010: 80: 186-192.

Greeff JW, Hartman E, Mullender-Wijnsma MJ, Bosker RJ, Doolaard S, Visscher C. Physical fitness and academic performance in primary school children with and without a social disadvantage. Health Educ Res 2014: 29: 853-860.

Grissom JB. Physical fitness and academic achievement. JEP 2005: 8: 11-25.

Gunnarsdottir T, Njardvik U, Olafsdottir AS, Craighead LW, Bjarnason R. Teasing and social rejection among obese children enrolling in family-based behavioural treatment: effects on psychological adjustment and academic competencies. Int J Obes 2011: 36: 35-44.

Haapala EA. Cardiorespiratory fitness and motor skills in relation to cognition and academic performance in children- A review. J Hum Kinet 2013: 36: 55-68.

Haapala EA, Poikkeus AM, Tompuri T, Kukkonen-Harjula K, Leppänen PH, Lindi V, Lakka TA. Associations of motor and cardiovascular performance with academic skills in children. Med Sci Sports Exerc 2014: 46: 1016-1024.

Hanna SL. The relationship between physical fitness and academic achievement in ninth-grade students in Arkansas. ProQuest LLC, Ed.D. Dissertation, Oral Roberts University 2009.

Hannigan WJ. The role of physical fitness in academic achievement. ProQuest LLC, Ed.D. Dissertation, Walden University 2010.

Hansen DM, Herrmann SD, Lambourne K, Lee J, Donnelly JE. Linear/Nonlinear relations of activity and fitness with children's academic achievement. Med Sci Sports Exerc 2014: 46: 2279-2285.

Hardy LL, Reinten-Reynolds T, Espinel P, Zask A, Okely AD. Prevalence and correlates of low fundamental movement skill competency in children. Pediatrics 2012: 130: 390-398.

Heberle AE, Carter AS. Cognitive aspects of young children's experience of economic disadvantage. Psychol Bull 2015: 141: 723-746.

Hillman CH, Castelli DM, Buck SM. Aerobic fitness in neurocognitive function in healthy preadolescent children. Med Sci Sports Exerc 2005: 37: 1967-1974.

Hillman CH, Erickson KI, Kramer AF. Be smart, exercise your heart: exercise effects on brain and cognition. Nat Rev Neurosc 2008: 9: 58-65.

Hillman CH, Buck SM, Themanson JR, Pontifex MB, Castelli DM. Aerobic fitness and cognitive development: Event-Related Brain Potential and task performance indices of executive control in preadolescent children. Dev Psychol 2009: 45: 114-129.

Huang T, Tarp J, Domazet SL, Thorsen AK, Froberg K, Andersen LB, Bugge A. Associations of adiposity and aerobic fitness with executive function and Math performance in Danish adolescents. J Pediatr 2015: 167: 810-815.

Huizinga MM, Beech BM, Cavanaugh KL, Elasy TA, Rothman RL. Low numeracy skills are associated with higher BMI. Obesity 2008: 16: 1966-1968.

Hupert J. Translating best evidence into best care. J Pediatr 2015: 166: 206-208.

Kantomaa MT, Purtsi J, Taanila AM, Remes J, Viholainen H, Rintala P, Ahonen T, Tammelin TH. Suspected motor problems and low preference for active play in childhood are associated with physical inactivity and low fitness in adolescence. PLoS One 2011: 6: e14554.

Kantomaa MT, Stamatakis E, Tammelin T. Physical activity and obesity mediate the association between childhood motor function and adolescents' academic achievement. Proc Natl Acad Sci 2013: 110: 1917-1922.

Kalantari HA, Esmaeilzadeh S. Association between academic achievement and physical status including physical activity, aerobic and muscular fitness tests in adolescent boys. Environ Health Prev Med 2016: 21: 27-33.

Keeley TJH, Fox KR. The impact of physical activity and fitness on academic achievement and cognitive performance in children. Int Rev Sport Exerc Psychol 2009: 2: 198-214.

Kim HP, Frongillo EA, Han S, Oh SH, Kim WK, Jang YA, Won HS, Lee HS, Kim SH. Academic performance of Korean children is associated with dietary behaviours and physical status. Asia Pac J Clin Nutr 2003: 12: 186-192.

Knaepen K, Goekint M, Heyman EM, Meeusen R. Neuroplasticity- Exercise-induced response of peripheral Brain-Derived Neurotrophic Factor. Sports Med 2010: 40: 765-801.

Korsten-Reck U, Kaspar T, Korsten K, Kromeyer-Hauschild K, Bös K, Berg A, Dickhuth HH. Motor abilities and aerobic fitness of obese children. Int J Sports Med 2007: 28: 762-767.

Kvaavik E, Klepp KI, Tell GS, Meyer HE, Batty GD. Physical fitness and physical activity at age 13 years as predictors of cardiovascular disease risk factors at ages 15, 25, 33 and 40 years: Extended follow-up of the Oslo Youth Study. Pediatrics 2009: 123: 80-86.

Kwak L, Kremers SP, Bergman P, Ruiz JR, Rizzo NS, Sjöström M. Associations between physical activity, fitness, and academic achievement. J Pediatr 2009: 155: 914-918.

Janak JC, Gabriel KP, Oluyomi AO, Peréz A, Kohl HW, Kelder SH. The association between physical fitness and academic achievement in Texas state house legislative districts: an ecologic study. J Sch Health 2014: 84: 533-542.

Lambourne K, Hansen DM, Szabo AN, Lee J, Herrmann SD, Donelly JE. Indirect and direct relations between aerobic fitness, physical activity, and academic achievement in elementary school students. Ment Health and Phys Act 2013: 6: 165-171.

Latorre PA, Mora LD, García PF. Association between intellectual maturity with physical fitness in preschool children. Pediatr Int 2015. Epub ahead of print.

Liao PA, Chang HH, Wang JH, Wu MC. Physical fitness and academic performance: Empirical evidence from the national administrative senior high school student data in Taiwan. Health Educ Res 2013: 28: 512-522.

London RA, Castrechini SA. A longitudinal examination of the link between youth physical fitness and academic achievement. J Sch Health 2011: 81: 400-408.

Lubans DR, Morgan PJ, Cliff DP, Bamett LM, Okely AD. Fundamental movement skills in children and adolescents: review of associated health benefits. Sports Med 2010: 40: 1019-1035.

Mobilia-Jones K. A study of the physical fitness test in relation to demographics, academic achievement, and students' physical fitness perceptions. ProQuest LLC, Ed.D. Dissertation, Pepperdine University 2010.

Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009: 151: 264-269.

Morita N, Nakajima T, Okita K, Ishihara T, Sagawa M, Yamatsu K. Relationship among fitness, obesity, screen time and academic achievement in Japanese adolescents. Physiol Behav 2016: 163: 161-166.

Nevill AM, Tsiotra G, Tsimeas P, Koutedakis Y. Allometric associations between body size, shape, and physical performance of Greek children. Pediatr Exerc Sci 2009: 21: 220-232.

Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. Int J Obes (Lond) 2008: 32: 1-11.

Padilla-Moledo C, Ruiz JR, Ortega FB, Mora J, Castro-Piñero J. Associations of muscular fitness with psychological positive health, health complaints, and health risk behaviors in Spanish children and adolescents. J Strength Cond Res 2012: 26: 167-173.

Pahkala K, Hernelahti M, Heinonen OJ, Raittinen P, Hakanen M, Lagström H, Viikari JS, Rönnemaa T, Raitakari OT, Simell O. Body mass index, fitness and physical activity from childhood through adolescence. Br J Sports Med 2013: 47: 71-77.

Pan L, Sherry B, Park S, Blanck HM. The association of obesity and school absenteeism attributed to illness or injury among adolescents in the United States. J Adolesc Health 2013: 52: 64-69.

Pellicer-Chenoll M, Garcia-Massó X, Morales J, Serra-Año P, Solana-Tramunt M, Gonzáles L-M, Toca-Herrera J-L. Physical activity, physical fitness and academic achievement in adolescents: a self-organizing maps approach. Health Educ Res 2015: 30: 436-448.

Pindus DM, Drollette ES, Scudder MR, Khan NA, Raine LB, Sherar LB, Esliger DW, Kramer AF, Hillman CH. Moderate-to-vigorous physical activity, indices of cognitive control, and academic achievement in preadolescents. J Pediatr 2016: 173: 136-142.

Raine LB, Lee HK, Saliba BJ, Chaddock-Heyman L, Hillman CH, Kramer AF. The influence of childhood aerobic fitness on learning and memory. Plos One 2013: 8: 1-6.

Rauner RR, Walters RW, Avery M, Wanser TJ. Evidence that aerobic fitness is more salient than weight status in predicting standardized math and reading outcomes in fourth-through eight-grade students. J Pediatr 2013: 163: 344-348.

Roberts CK, Freed B, McCarthy WJ. Low Aerobic Fitness and Obesity Associated with Lower Standardized Test Scores in Children. J Pediatr 2010: 156: 711-718.

Roebers CM, Kauer M. Motor and cognitive control in a normative sample of 7-year-olds. Dev Sci 2009: 12: 175-181.

Rowland T, Kline G, Goff D, Martel L, Ferrone L. One-mile run performance and cardiovascular fitness in children. Arch Pediatr Adolesc Med 1999: 153: 845-849.

Rowland T. Oxygen uptake and endurance fitness in children, revisited. Pediatr Exerc Sci 2013: 25: 508-14.

Ruiz JR, Ortega FB, Gutierrez AM, Sjöström DM, Castillo MJ. Health-related fitness assessment in childhood and adolescence: A European approach based on the AVENA, EYHS and HELENA studies. J Public Health 2006: 14: 269-277.

Ruiz JR, Castro-Pinero J, Artero EG, Ortega FB, Sjöström M, Suni J, Castillo MJ. Predictive validity of health-related fitness in youth: a systematic review. Br J Sports Med 2009: 43: 909-23.

Ruiz JR, Castro-Piñero J, España-Romero V, Artero EG, Ortega FB, Cuenca MM, Jimenez-Pavón D, Chillón P, Girela-Rejón MJ, Mora J, Gutiérrez A, Suni J, Sjöström M, Castillo MJ. Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. Br J Sports Med 2011: 45: 518-24.

Saavedra JM, Torres S, Caro B, Escalante Y, De La Cruz E, Durán MJ, Rodríguez FA. Relationship between health-related fitness and educational and income levels in Spanish women. Public Health 2008: 122: 794-800.

Sardinha LB, Marques A, Martins S, Palmeira A, Minderico C. Fitness, fatness, and academic performance in seventh-grade elementary school students. BMC Pediatrics 2014: 14: 1-9.

Sardinha LB, Marques A, Minderico C, Palmeira A, Martins S, Santos DA, Ekelund U. Longitudinal Relationship between Cardiorespiratory Fitness and Academic Achievement. Med Sci Sports Exerc 2016: 48: 839-844.

Scudder MR, Federmeier KD, Raine LB, Direito A, Boyd JK. The association between aerobic fitness and language processing in children: implications for academic achievement. Brain Cogn 2014: 87: 150-152.

Schwimmer JB, Burwinkle TM, Varni JW. Health-related quality of life of severely obese children and adolescents. JAMA 2003: 289: 1813-1819.

Shore SM, Sachs ML, Lidicker JR, Brett SN, Wright AR, Libonati JR. Decreased scholastic achievement in overweight middle school students. Obesity 2008: 16: 1535-1538.

Srikanth S, Petrie TA, Greenleaf C, Martin SB. The relationship of physical fitness, self-beliefs, and social support to the academic performance of middle school boys and girls. J Early Adolescence 2014: 35: 1-25.

Telford RD, Cunningham RB, Telford RM, Abhayaratna WP. Schools with fitter children achieve better literacy and numeracy results: evidence of a school cultural effect. Pediatr Exerc Sci 2012: 24: 45-57.

The IDEFICS study, Zagout M, Michels N, Bammann K, Ahrens W, Spregeler O, Molnar D, Hadjigeorgiou C, Eiben G, Konstabel K, Russo P, Jimenez D, Moreno LA, De Henauw S. Influence of physical fitness on cardio-metabolic risk factors in European children. Int J Obes (London) 2016. Epub ahead of print.

Tokmakidis SP, Kasambalis A, Christodoulus AD. Fitness levels of Greek primary schoolchildren in relationship to overweight and obesity. Eur J Pediatr 2006: 165: 867-874.

Torrijos-Niño C, Martínez-Vizcaíno V, Pardo-Guijarro MJ, García-Prieto JC, Arias-Palencia NM, Sánchez-López M. Physical fitness, obesity, and academic achievement in schoolchildren. J Pediatr 2014: 165: 104-109.

Tucker-Drob EM, Harden KP. Early childhood cognitive development and parental cognitive stimulation: evidence for reciprocal gene-environment transactions. Dev Sci 2012: 15: 250-259.

Van der Fels IM, Te Wierike SC, Hartman E, Elferink-Gemser MT, Smith J, Visscher C. The relationship between motor skills and cognitive skills in 4-16 year old typically developing children: A systematic review. J Sci Med Sport 2015: 18: 697-703.

Van Dusen DP, Kelder SH, Kohl HW, Ranjit N, Perry CL. Associations of physical fitness and academic performance among schoolchildren. J Sch Health 2011: 81: 733-740.

Van Der Niet AG, Hartman E, Smith J, Visscher C. Modeling relationship between physical fitness, executive functioning, and academic achievement in primary school children. Psychol Sport Exerc 2014: 15: 319-325.

Valkeinen H, Harald K, Borodulin K, Mäkinen TE, Heliövaara M, Leino-Arjas P, Sainio P, Kestilä L, Kunst A, Rahkonen O, Tammelin T, Härkänen T, Prättälä R. Educational differences in estimated and measured physical fitness. Eur J Public Health 2013: 23: 998-1002.

Welk GJ, Jackson AW, Morrow JR, Haskell WH, Meredith MD, Cooper KH. The association of health-related fitness with indicators of academic performance in Texas schools. Res Q Exerc Sport 2010: 81: 16-23.

Wingfield RJ, McNamara JPH, Janicke DM. Is there a relationship between Body Mass Index, fitness, and academic performance? Mixed results from students in a Southeastern Unites States elementary school. Curr Issues Educ 2011: 14: 1-12.

Wittberg RA, Northrup KL, Cottrel L. Children´s physical fitness and academic performance. Am J Health Educ 2009: 40: 30-36.

Wittberg RA, Cottrell LA, Davis CL, Northrup KL. Aerobic fitness thresholds associated with fifth grade academic achievement. Am J Health Educ 2010: 41: 284-291.

Wittberg RA, Northrup KL, Cottrell LA. Children's aerobic fitness and academic achievement: A longitudinal examination of students during their fifth and seventh grade years. Am J Public Health 2012: 102: 2303-2307.

Fig. 1. Flow chart of the study selection process..

Table 1. Main results of studies about Physical fitness and academic performance in children and adolescents.

References	Study design and location	Participants characteristics	Assessment of PF	Assessment of academic performance	Control variables	Main results
Du Toit et al. (2011)	Cross-sectional/ South Africa	212 students/ 9 to 12 age/ 4th to 5th grades	Battery of physical tests of the Fitnessgram (body composition, CRF-PACER, muscular strength, muscular endurance and flexibility) and Bruininks-Oseretsky Test of Motor Proficiency II.	The average of the end-of-the-year academic marks, as recorded in school schedules and children`s report cards according to the prescriptions of the National Department of Education for learners in the Intermediate Phase (Grade 4, 5 and 6), was used as a measure of academic achievement.	No utilized	Significant positive correlation between total strength scores and AP in girls (r=0.35; p=,0.05). A positive relationship between physical fitness components and academic achievement was found with more significant correlations among girls (CRF r=0.63; p=<0.05, knee push-ups r=0.62; p<0.05, wall sitting r=0.54; p=<0.05; strength total r=0.62; p=<0.05) than boys (CRF r=0.40; p=>0.05, knee push-ups r=0.45; p>0.05, wall sitting r=0.16; p=>0.05; strength total r=0.51; p=<0.05), as well as among older boys and girls.
Kim et al. (2003)	Cross-sectional/ South Korea	6,463 students/ 10 to 14 age/ 5th to 8th grades	BMI, Sprint (100m) and long (600-1000m) distance race, standing broad jump, back up exercise throwing, and chin-ups for boys or hanging on the bar for girls.	Reading, math, social studies, science, physical education, music, arts and ethics.	Parental education level was used to represent the SES. Dietary behaviour and food frequency was collected.	Physical fitness score was positively associated with AP in boys (β=0.101; p=<0.001) and girls (β=0.050; p=<0.01). There was no association between AP and body mass.
Eveland-Sayers et al. (2009)	Cross-sectional/ United States	134 students/ 9 to 11 age/ 3th to 5th grades	BMI, Test of a mile, abdominal and flexibility.	Terra Nova achievement test (assessment in reading, language and mathematics)	No utilized	There was a statistically significant negative relationship (r=-0.31; p=<0.001 and r=-0.36; p=<0.05, respectively) among 1-mile run times and reading/language arts and mathematics test scores for girls. There was no relationship

						between BMI and academic performance.
Wittberg et al. (2009)	Cross-sectional/ United States	741 students/ 9 to 13 age/ 5th grade	Battery of physical tests of the Fitnessgram (body composition, CRF-PACER or one mile run, muscular strength, muscular endurance and flexibility).	West Virginia Educational Standards Test (WESTEST) which assesses language, reading, math, science and social studies.	Meal program status (whether the child received a free lunch, a reduced cost lunch, or paid for their lunch) served as a proxy for SES.	All children who were in the Healthy Fitness Zone (HFZ) for CRF and abdominal strength scored significantly higher on reading (3.34; 3.31), math (3.51; 3.46), Science (3.47; 3.43) and social studies (3.40; 3.38) than those children who were in the Needs Improvement Zone (NIZ) (3.09; 2.97 reading), (3.14; 2.98 math), (3.25; 3.20 science) and (3.20; 3.07). Math scores were significantly higher among children who were in the HFZ for the upper body strength test (3.44) and flexibility test (3.43) than children in the NIZ (3.26; 3.13). Science scores were also significantly higher for students in the HFZ for flexibility (3.43) than students in the NIZ (3.20). There was no association between BMI and academic performance.
Roberts et al. (2010)	Cross-sectional/ United States	1,989 students/ 10 to 16 age/ 5th to 9th grades	Test of a mile	Achievement Tests version e o California Standards Test (CST), using math and language.	Parental education, child ethnicity, and eligibility for free or reduced-price lunch status were determined by parent self-report information collected by the school district.	Mile run/walk time was a significant predictor of performance in math such that the math score dropped 1.9 points for every additional minute required to complete the 1-mile run/walk (β=-1.94; CI=-2.37, -1.53). Reading test score dropped 1.1 points for every additional minute required to complete the 1-mile run/walk (β= -1.13; CI=-1.56, -0.70). Negative associations between overweight and academic performance.
Davis & Cooper (2011)	Cross-sectional/ United States	170 students/ 7 to 11 age/ 1st to 5th grades	Incremental treadmill test	Woodcock-Johnson test (assessment in reading and math) and cognitive skills.	Race, gender, and primary caregiver's education level.	CRF was positively associated with cognition (planning r=0.26; p=<0.001, attention r= 0.22; p=<0.001), performance in reading (r=0.29; p=<0.001) and math (r=0.25; p=<0.001). Body fat (planning r=-0.22; p=<0.001, attention r=-0.16; p=<0.05; math r=-0.23; p=<0.001; reading r=-

						0.20; p=<0.05) and WC were negatively related (planning r=-0.16; p=<0.05; math r=-0.28; p=<0.001; reading r=-0.21; p=<0.05.
London & Castrechini (2011)	Longitudinal/ prospective cohort/ United States	1,410 students/ 9 to 15 age/ 4th to 9th grades	Battery of physical tests of the Fitnessgram (body composition, CRF- one mile run, muscular strength, muscular endurance and flexibility).	California standardized test (CST), assessment in reading and math.	School context, SES, ethnicity, proximity to fast food restaurants, and opportunities for PA.	For the younger cohort, students who go on to pass in physical fitness tests and students who fail in the first year but pass in the second have higher math scores in the fourth grade compared to students who go on to fail both tests (β=0.244; p=<0.001 and β=0.147; p=<0.01, respectively). Similarly, both these 2 groups of students have significantly higher scores on the reading than those who fail both tests (β=0.115; p=<0.05 and β=0.129; p=<0.05, respectively). For the older cohort, an initial CST gap is present between those passing and failing both tests (β=0.218; p=<0.001 in math, β=0.154; p=<0.01 in reading). Models that focus on the most challenging of the tests-the mile run and the push-up test and find no conclusive evidence that one of these tests, rather than the combination of all of them, is responsible for the link with academic achievement. There were no associations between BMI and academic performance.
Van Dusen et al. (2011)	Cross-sectional/ United States	254,743 students/ 9 to 16 age/ 3th to 11th grades	Battery of physical tests of the Fitnessgram (body composition, CRF-PACER or one mile run, muscular strength, muscular endurance and flexibility).	Texas Assessment of Knowledge and Skills (TAKS), (assessment in reading and math).	Gender, age, grade ethnicity, economic disadvantage (measure by school lunch status), school number, and special educational status.	CRF was found to have the strongest direct associations with academic achievement, with a standardized mean difference effect size of 0.17 (CI:0.15-0.19) for boys-reading, 0.34 (CI:0.32-0.35) for boys-math, 0.27 (CI:0.25-0.29) for girls reading, 0.33 (CI:0.31-0.35) for girls-math. The next largest associations were with curl-ups, followed by push-ups, sit and reach, and trunklift which registered the lowest effect size of .07 (CI:0.05-0.08).

Wingfield et al. (2011)	Cross-sectional/ United States	132 students/ 10 years old/ 4th to 5th grades	BMI, Curl-ups, 20-meter shuttle run test, one mile run, pull-ups, flexed arm hang, and V-sit reach.	Florida Comprehensive Assessment Tests (FCAT), to asses reading and math.	Sex, age, race, and SES.	In girls, only BMI marginally predicted AP (β=-0.40; p=<0.05), fitness no longer significantly predicted AP when BMI was in the model (β=0.21; p=0.34).
Chen et al. (2013)	Longitudinal/ prospective cohort/ Taiwan	669 students/ 13 to 15 age/ 7th to 9th grades	BMI, 1,600-Meter run (boys), 800-meter run (girls), sit-and-reach test, curl-ups.	Mean score of Language, Mathematics, Science, and Social Studies.	Measures of SES, parental education and ethnicity.	CRF exhibits stronger longitudinal associations with AP than other forms of fitness or BMI for adolescents. CRF slope to the academic slope was significant, showing the rate of change in CRF was related to the rate of change in AP over this period (z=6.98; p=<0.05).
Coe et al. (2013)	Cross-sectional/ United States	1,701 students/ No age available/ 3rd, 6th, 9th grades	Battery of physical tests of the Fitnessgram (BMI, CRF-PACER, muscular strength, muscular endurance and flexibility).	Michigan Education Assessment Program (MEAP) that asses English/language mathematics and social studies.	Student eligibility for a free and reduced lunch program was used as a proxy for SES, based on household income and number of family members in the home.	In 6th, Muscular strength and endurance were significantly associated with performance in math (r=0.149; p=<0.001), English (0.106; p=<0.001) and social studies (r=0.144; p=<0.001). However, compared with all other variables, SES had the strongest association with academic achievement in 6th and 9th, respectively (r=0.183-0.324; p=< 0.05). Regression analysis yielded similar results, with SES accounting for a large portion of the variance in standardized test percentiles (18.6-32.2%; p=<0.001). This was true for all subject areas in all grades, except for Math in ninth-grade students. In this group, curl-ups were the best predictor of academic achievement (accounting for 20.9% of the variance; p =< 0.001), but SES was also significant in this model (accounting for 6.2% of the variance; p=< 0.001).
Lambourne et al. (2013)	Cross-sectional/ United States	401 students/ 6 to 9 age/ 2nd to 3rd grades	20-meter shuttle run test, accelerometer to assess daily PA.	Weschsler Individual Achievement Test-Third Edition.	Grade, age, gender, ethnicity, and race, as well as the education level of the mother and	PA had a direct effect on CRF (0.009; p=<0.01), which in turn had a direct effect on math scores (-0.007; p=<0.05). The effect of PA on math achievement was partially mediated by CRF (0.003; p;<0.05).

					household income. Daily PA.	
Haapala et al. (2014)	Longitudinal/ prospective/ Finland	167 students/ 6 to 8 age/ 1st to 3th grades	Maximal exercise test (cycle ergometer). Shuttle run test was used to assess speed and agility, the flamingo balance test was used to assess static balance and box and block test was used to assess manual dexterity.	Reading fluency and comprehension were assessed using a group-administered speeded subtest of the nationally-normed reading achievement battery. Arithmetic skills were assessed using a basic arithmetic test with a set of visually presented addition and subtraction tasks.	Measures of parental education, PA level, pubertal status and risk of reading disabilities.	Among boys, overall motor performance in grade 1 was positively associated with reading fluency in grades 1–3 (β=0.40; p=<0.001, β=-0.41; p=<0.001), reading comprehension in grades 1 and 2 (β=0.39; p=<0.001, β=0.25; p=<0.05), and arithmetic skills in grades 1–3 (β=0.41; p=<0.001, β=0.39, p=<0.001) after adjustment for age, parental education, and other measures of cardiovascular and motor performance. Maximal workload in grade 1 was not associated with academic skills in grades 1–3.
Srikanth et al. (2014)	Cross-sectional/ United States	1,211 students/ 12 to 13 age/ 6th to 8th grades	Progressive Aerobic Cardiovascular Endurance Run (PACER)	Texas Assessment of Knowledge and Skills (TAKS), (assessment in reading and math).	SES was based on federal guidelines for determining students' status for free or reduced lunch based on family income. PA and psychosocial variables, general self-concept and social support.	Boys being more aerobically fit were associated with better performances in the TAKS reading exam (β=0.24; p=<0.005) and math (β=0.26; p=<0.005). Girls more aerobically fit were associated with better performances in the TAKS reading exam (β=0.23; p=<0.005) and math (β=0.17; p=<0.005).
Pindus et al. (2016)	Cross-sectional/ United States	74 students/ 7 to 9 age/ No data grades available	Graded treadmill test using a computerized indirect calorimetry system.	Kaufman Test of Educational Achievement, Second Edition (assessment in reading and math) and standardized score for spelling subscale were included.	Age, gender, Intelligence quotient, PA, SES and clinical diagnosis of attention deficit hyperactivity disorder.	CRF was positively associated with spelling (β=-0.22; p= 0.04; F(2, 71) = 12.8, P < .001) after adjusted by covariates.

Castelli et al. (2007)	Cross-sectional/ United States	259 students/ 9 to 11 age/ 3th to 5th grades	Battery of physical tests of the Fitnessgram (BMI, CRF-PACER, muscular strength, muscular endurance and flexibility). .	Illinois Standards Achievement Test (ISAT) (assessment in reading and math).	Age, gender and SES measures.	There were significant effects for BMI and for CRF (β=−0.15; p=< 0.01), (β=0.40; p=< 0.001) indicating that lower BMI and higher CRF were positively related to reading achievement and math (β=-0.13; p=<0.05), (β=0.42; p=<0.001).
Chomitz et al. (2009)	Cross-sectional/ England	2,217 students/ 10 to 14 age/ 4th to 8th grades	Physical tests adapted from the Amateur Athletic Union and Fitnessgram (BMI, CRF test, abdominal strength test, flexibility, upper body strength test, and agility test).	MCAS tests (assessment in reading and math).	Gender, SES, weight status, grade, and ethnicity.	The logistic regression analysis estimated that the odds of passing the Math MCAS increased by 38% for each 1-unit increase in the number of fitness tests passed, holding gender, ethnicity, weight status, grade, and SES constant (OR=1.379; CI=1.234-1.541) the odds of passing the reading MCAS test increased by 24% for each 1-unit increase in the number of fitness tests passed (OR=1.236; CI=1.003-1.522).
Wittberg et al. (2010)	Cross-sectional/ United States	1,740 students/ 9 to 13 age/ 5th grade	Test of a mile and Progressive Aerobic Cardiovascular Endurance Run (PACER).	West Virginia Educational Standards Test (WESTEST) which assesses language, reading, math, science and social studies.	No utilized	In boys, faster students in mile test had higher average in reading (3.75; p=<0.001), math (3.37; p=<0.001), science (3.69; p=<0.01) and social studies (3.48; p=<0.001) when compared with slower students, respectively, (3.26; 2.91; 3.39; 3.12). The same was not found when the CRF was measured by shuttle run test. In girls, faster student in the PACER test had higher average in reading (3.43; p=<0.001), math (3.37; p=<0.001) and social studies (3.22; p=<0.01) when compared with slower students (3.18; 3.20; 3.16). However, when CRF was measured by the test of a mile there were no associations.
Blom et al. (2011)	Cross-sectional/ United States	2,992 students/ No age reported/	Battery of physical tests of the Fitnessgram (BMI, CRF-PACER, muscular strength,	Mississippi Curriculum Test (MCT2) that assesses performance in language arts	Academic behavior data on absences and disciplinary	A linear trend was observed between language arts, Math and overall fitness level, indicating the odds of high academic achievement increased

		3th to 8th grades	muscular endurance and flexibility).	and mathematics.	actions, and socio-demographic information of gender, race/ethnicity, and SES (via lunch status), gender and race.	with the number of healthy fitness zones achieved (OR=3.31; CI=1.66-6.57), (OR=4.14; CI=2.01-8.55), respectively. The relationships remained significant while controlling for gender, race, and SES.
Wittberg et al. (2012)	Longitudinal/ prospective cohort/ United States	1,725 students/ No date age available/ 5th to 7th grades	20-meter shuttle run test	West Virginia Educational Standards Test (WESTEST) which assesses language, reading, math, science and social studies.	No utilized	Students at the beginning of the cohort had CRF in Healthy Fitness Zone and those during follow-up can achieve it have higher AP (0.133; p=<0.001, 0.127; p=<0.01) than students who started outside of this zone (-0276; -0.284).
Bass et al. (2013)	Cross-sectional/ United States	838 students/ 13 years old/ 6th to 8th grades	Battery of physical tests of the Fitnessgram (BMI, body composition, CRF-PACER, muscular strength, muscular endurance and flexibility).	Illinois Standards Achievement Test (ISAT) which assesses math and reading	Age, sex, participation in the National School Lunch Programme, providing free or reduced lunch to economically disadvantaged students, was used as an indicator for socio-economic status.	Boys in the Healthy Fitness Zone (HFZ) for CRF or muscular endurance had higher odds in Reading performance (3.15; CI= 1.90-5.21), (2.56; CI=1.22-5.38) than boys not in the HFZ. Girls in the HFZ for CRF had higher odds in reading and math test performance (2.08; CI=1.10-3.93, 2.58; CI=1.39-4.78) than girls not in the HFZ, respectively. There were no associations between body composition and academic performance.
Kantomaa et al. (2013)	Longitudinal prospective/ Finland	8,061 students/ 7 to 16 age follow-up/ No data grades available	Submaximal cycle ergometer test and gross motor skills at the age 8 y through the parents' questionnaire.	The GPA was calculated as a measure of academic achievement based on the grades in the following school subjects: mother tongue, first foreign language (started at grade 3), second foreign language (started at	Mother's education, BMI, PA and motor function.	Compromised motor function in childhood had a negative indirect effect on academic achievement via lower levels of PA (β= –0.023; p=<0.001) and obesity (β=-0.025; p=<0.001), but not via CRF (β=0.004; p=0.314).

				grade 7), math, biology, geography, physics, chemistry, religion or ethics, history, music, visual arts, physical education, crafts, and home economics.		
Liao et al. (2013)	Longitudinal/ prospective/ Taiwan	149,240 students/ 17 to 18 age/ Senior high school	1,600-meter run (boys), 800-meter run (girls), sit-and-reach test, curl-ups.	General Scholastic Ability Test (GSAT) that asses chinese, english, mathematics, science and social studies.	SES and age.	On average, 1 Standard Deviation (SD) increase on the anaerobic power and flexibility Z-scores from the first to third year of senior high school in girls are associated with an increase in the university entrance exam by 0.018 and 0.010 SD, respectively.
Rauner et al. (2013)	Cross-sectional/ United States	11,742 students/ No date age available/ 3th to 8th grades	20-meter shuttle run test	Nebraska State Accountability (NESA) which assesses math and reading.	BMI percentile, SES, sex, ethnicity, grade level, and school type.	Aerobically fit students receiving free/reduced lunch had 1.68 times greater odds of passing the NeSA math test and 1.56 times greater odds of passing the NeSA reading test compared with aerobically unfit students receiving free/reduced lunch. BMI was not associated with academic performance.
Greeff et al. (2014)	Cross-sectional/ Netherlands	544 students/ 7 to 10 age/ 2nd to 3rd grades	Physical fitness was evaluated with items of the Eurofit physical fitness test battery (BMI, CRF, explosive strength, abdominal muscle endurance and static strength)	AP was evaluated with scores on mathematics and two domains of language, namely spelling and reading.	Age, grade, gender and SES.	Results showed a positive association between CRF and mathematics (β=0.23; p=<0.01), regardless of SES. A positive association was found between CRF and spelling performance (β=0.16; p=<0.05). However, lower SES did moderate the relationship between CRF and spelling.
Hansen et al. (2014)	Cross-sectional/ United States	687 students/ 7 to 8 age/ 2nd to 3rd grades	Progressive Aerobic Cardiovascular Endurance Run (PACER)	Weschsler Individual Achievement Test-Third Edition (WIAT-III) was used to assess academic achievement in reading,	PA, gender, race, ethnicity, mother's education, SES and BMI.	No significant linear or non-linear association between CRF and reading achievement (β=0.06; p=0.46), however, a significant quadratic association between CRF and spelling (β=0.32; p=<0.01) and math achievement (β=0.55;

				spelling and math.		p=<0.01).
Scudder et al. (2014)	Cross-sectional/ United States	46 students/ 9 to 10 age/ No data grades available	Computerized indirect calorimetry system.	Reading, spelling, and arithmetic achievement were determined using the WRAT3. Hand dominance, intelligence quotient and academic achievement were also determined using Edinburgh Handedness Inventory and Kaufman Brief Intelligence Test 2 (K-BIT2). Additionally, measures of Event-related brain potentials (ERPs) were realized through electroencephalographic (EEG) and eletrooculographic (EOG)	SES and pubertal stage.	Higher fit children had greater reading achievement scores (123.1 ± 2.8) compared to lower fit children (112.9 ± 2.1), t (44)= 2.8, p= 0.007, d = 0.8. A similar finding was observed for spelling (higher fit: 117.6 ± 2.6; lower fit: 108.0 ± 3.0), t(44)= 2.4, p= 0.02, d = 0.7. Higher fit children have faster neuroelectric responses and perform better on cognitive tests including those linked to language skills.
Torrijos-Niño et al. 2014	Cross-sectional/Spain	893 students/ 9 to 10 age/ 4th to 5th grades	Alpha Battery tests (BMI, CRF, muscular fitness handgrip test, speed/agility and standing broad jump test)	Final grades of the participants the previous year (2009/2010). Scores obtained in Math, language and literature, natural, social and cultural sciences, and English.	Age, sex, and parental education.	Boys with good CRF and speed/agility had, after controlling for potential confounders, showed 7.3 and 4.0 times higher probability of score in the top quartile of academic achievement than children with poor CRF and speed/agility levels. Girls with satisfactory or good speed/agility had a higher probability of scoring in the top quartile of academic achievement (3.8 and 4.7 times, respectively) than girls with poor speed/agility after controlling for confounders. Physical fitness is more closely related to academic achievement than BMI.
Chaddock-Heyman et al. (2015)	Cross-sectional/United States	42 students/ 9 to 10 age/ No data grades available	Computerized indirect calorimetry system.	Wide Range Achievement Test-3rd edition (WRAT-3) was used to assess academic achievement in reading,	Pubertal timing and SES.	Higher fit children showed decreased cortical thickness in superior frontal cortex (F (1, 46) =4.80, p = 0.034), superior temporal cortex (F (1,

				spelling and math. T1-weighted structural brain images were acquired using a 3D MPRAGE (Magnetization Prepared Rapid Gradient Echo).		46) = 5.39, p = 0.025) and lateral occipital cortex (F (1, 46) = 5.67, p = 0.021), relative to lower fit children. Higher fit children showed superior mathematics achievement compared to lower fit children (t (46) = 1.98, p = 0.05) on the WRAT-3. No fitness differences were found for reading or spelling performance (t< 1.1, p>0.3).
Desai et al. (2015)	Cross-sectional/India	273 students/7 to 10 age/ No data grades available	BMI, body composition and 20-m shuttle run test.	School-wide exams administered by school-teachers. Subject areas assessed included mathematics, Kannada (the official language of the Indian state of Karnataka), English, science, and social science. Only math and Kannada scores, reported as percentages, were used as final academic outcome measures because these two subjects were assessed in all participants.	Physical activity, diet, age, SES and micronutrient status.	After standardizing scores across grade levels and adjusting for school, gender, SES, and BMI, children with greater CRF had greater odds of scoring above average on math and Kannada exams (OR=1.08; CI: 1.02-1.15 and OR=1.11; CI: 1.04-1.18, respectively).
Pellicer-Chenoll et al. (2015)	Longitudinal/ Spain	700 students enrolled in the first year of secondary education/ No age reported	BMI, Jump power was measured from the data obtained in a countermovement jump (CMJ), CRF was measured using the Cooper test, handgrip isometric strength was measured with a Takei dynamometer.	Academic achievement was obtained from the student's transcripts using the GPAs. The procedure was as follows: the mean value of 10 academic topics corresponding to an academic year (GPA) was obtained for each of the 4 years that the students spent at secondary school. The score of the mean value ranged from 0 (worst rating)	No utilized	Students with lower energy expenditure (~1000 METs week^{-1}) and lower physical fitness exhibited higher BMI and low AP below 5 out of 10 points. whereas those adolescents with higher energy expenditure exhibited better physical fitness, lower BMI and higher AP.

				to 10 (best rating).		
García-Hermoso (2016)	Cross-sectional/ Chile	395 students/ aged 12/ 7th grades	BMI, Alpha Battery tests (20 m shuttle run test and standing broad jump test).	Students' grades in the core subjects (mathematics and language) were collected from the official school records at four moments in the first semester (March, April, May and June 2014). Grade scores in Chile range from 1 (worst) to 7 (best).	Physical activity, sedentary behavior, mother's education and SES.	Higher birth weight was associated with better AP independent of current BMI. These differences disappeared after controlling for CRF, which also mediated the association between birth weight and AP (13.4%; z=2.33; p=0.012). The relationship between birth weight and AP was not mediated by muscular strength.
Morita et al. (2016)	Cross-sectional/ Japan	315 students/ 12 to 13 age/ 7th grades.	BMI, 50-m sprint, standing broad jump, repeated side-steps, sit and reach, sit-ups, hand grip strength, handball throw, and 20-m shuttle run.	Academic achievement was assessed by the total grade point (GP) of school subjects evaluated by school teachers; individual grades were reported for 8 school subjects (Japanese, Mathematics, Social Studies, Sciences, English, Music, Arts, and Home Economics/Vocational Technology).	BMI, SES, mother's education and time spent on electronic devices.	After adjusting for covariates, total physical fitness was positively associated with AP in boys (β=0.13; p=<0.001). In girls, obesity status, but not physical fitness, was significantly associated with AP after adjusting for the confounders (β=-1.61; p=0.04).
Dwyer et al. (2001)	Cross-sectional/ Australia	9,000 students/ 7 to 15 age/ 1st a 9th grades	BMI, Long jump, abdominal, push-ups, flexibility, handgrip, cardiopulmonary exercise test and sprint 50 yards.	Evaluation overall accomplished through scale: excellent, above average, average, below average or poor.	Height, BMI, body fat, SES, time to be and length of sleep.	There were weak but consistent associations between field tests of muscular force, endurance and power. Students with higher AP took less time to complete the 50-meter run (r=-0.15; p=<0.001), completed more sit-ups (r=0.14; p=<0.001), and leapt greater distances in the standing long jump (r=0.10; p=<0.001). BMI was negatively associated with AP.
Kwak et al. (2009)	Cross-sectional/ Sweden	232 students/ 15 to 16 age/ 9th grade	Incremental test on a cycle ergometer and accelerometer 4 consecutive days	17 school subjects.	Mother's education, family structure, parental monitoring, sex, pubertal phase, PA,	Vigorous PA was the only intensity level that significantly correlated with academic achievement (β=023; p=<0.05), solely in girls, whereas in boys only fitness was associated positively with academic achievement (β=0.25;

					and sum of skinfolds.	p=<0.05).
Welk et al. (2010)	Cross-sectional/ United States	36,835 students/ 9 to 16 age/ 3th to 11th grades	20-meter shuttle run test	Texas Assessment of Knowledge and Skills (TAKS), (assessment in reading and math).	Grade level (elementary, middle, secondary), school size (number of students), gender, minority status, and SES.	Positive association between CRF and performance in TASK (OR=1.014; CI=1.011-1.016), indicating that for each 1% of higher CRF increase AP.
Telford et al. (2012)	Cross-sectional/ Australia	757 students/ 8 to 10 age/ 3th to 5th grades	20-meter shuttle run test, pedometers to assess daily PA.	Texas Assessment of Knowledge and Skills (TAKS), (assessment in reading and math).	PA, body fat % and SES.	School reading scores were positively associated with both the school CRF (β=0.143; p=< 0.001) and PA (β=4.7; p=0.01) and there was evidence that reading score was negatively related to % body fat at the school level (β=-5.5; p=0.05). School numeracy scores were positively associated with school CRF levels (β=0.142; p=<0.001) and PA (β=4.11; p=0.02) and associated negatively with % body fat (β=-6.5; p=0.02). School writing scores were associated with the school CRF (β=0.145; p= < 0.001) with little evidence of any association with PA (β=1.5; p=0.17) or % body fat (β=-3.0; p=0.23).
Bezold et al. (2014)	Longitudinal/ United States	83,111 students/ No data age available/ 6th to 8th grades	Battery of physical tests of the Fitnessgram (BMI, CRF-PACER, muscular strength and muscular endurance).	Measure of academic test performance, based on New York State standardized assessment in English Language Arts and Math.	Race and/or ethnicity, language spoken at home, days absent from school, place of birth, student household poverty, school-area poverty, and obesity status.	Girls who experienced an increase in fitness improved their academic test scores by 0.36 percentile points more per year than girls who experienced no change in fitness (β=0.36; CI:0.09-0.63). Substantial decreases in fitness were associated with a decline of 0.40 percentile points per year compared with the reference group (β=-0.40; CI=-0.68,-0.12), whereas moderate decreases in fitness were associated with a trend toward a decline in academic ranking

						of 0.33 percentile points per year (β=-0.33; CI=-0.67,-0.01) compared with the reference group. In boys, increasing fitness resulted in an improvement in academic ranking of 0.38 percentile points per year compared with the reference group (β=0.38; CI=0.09-0.66). A substantial decrease in fitness, academic ranking declined over time by 0.55 percentile points per year compared with the reference group (β=-0.55; CI=-0.85,-2.25).
Esteban-Cornejo et al. (2014)	Cross-sectional/ Spain	2,038 students/ 6 to 18 age/ No data grades available	Physical fitness was assessed following the ALPHA (Assessing Levels of PA) health-related fitness test battery for youths (BMI, body composition, Muscular strength, motor ability and CRF.	Four indicators were used to define AP: individual grades for the core subjects (math and language), an average for math and language, and Grade Point Average (GPA) score.	Sex, age, city, maternal education, pubertal status and waist circumference.	CRF and motor ability, both independently and combined, were related positively to AP in youths, independent of potential confounders, including fatness (β=0.136; p=<0.001. Muscular strength was not associated with AP independent of the other 2 physical fitness components (β=-0.005; p=0.882).
Janak et al. (2014)	Cross-sectional/ecologic study/ United States	2,550.114 students/ No data age available/ 3rd to 12th grades	Progressive Aerobic Cardiovascular Endurance Run (PACER) and the 1-mile run.	Texas Essential Knowledge and Skills curriculum (TAKS) that assesses performance in English language arts, reading, writing, mathematics, science, and social studies.	SES and grade category stratified by sex.	The magnitude of effect for CRF (β=0.56; p=< 0.001), BMI (β=1.14; p=< 0.001), and SES (β=−0.30; p=<0.001) in the univariate models were reduced by 61% (β=0.13; p=0.02), 77% (β=0.45; p=0.001), and 43% (β=−0.17; p=< 0.001), respectively in the multivariable model. SES explained the largest percent of the variance in academic achievement with BMI having the largest magnitude of effect.
Sardinha et al (2014)	Longitudinal/ prospective cohort/ Portugal	1,531 students/ 12 to 14 age/ 3rd to 9th grades	Aerobic Cardiovascular Endurance Run (PACER).	AP was assessed using the marks students had, at the end of their academic year, in math, language (Portuguese), foreign language (English) and	Gender, weight status, and different cohorts.	CRF and weight status were independently and combined related to academic achievement independent of the different cohorts. The odds of having high or average academic achievement versus low academic achievement of normal weight students was higher than overweight in

				sciences. Student marks range from 1 to 5 (very poor to very good).		both unadjusted (OR=4.98; CI:2.53-9.81) and adjusted analysis (OR=3.72; CI:1.85-7.49). Fit students, compared with unfit students, had higher odds for having a high academic achievement, in both the unadjusted (OR=2.27; CI:1.57-3.26), and adjusted model (OR=2.27; CI:1.46-3.52).
Aguilar et al. (2015)	Cross-sectional/ Chile	395 students/ 12 years old/ 7th grades	BMI, 20-meter shuttle run test, muscular strength was measured with the standing broad jump test. Questionnaire employed to assess PA	Academic attainment was assessed using the students' grades in the core subjects (mathematics and language).	Mother's and father's education, SES and screen time.	CRF was related to AP in both genders independent of potential confounders. However, these associations did not remain significant after adjusting for screen time. CRF was positively associated with language (β= 0.272, CI=0.658-2.365, β= 0.153; CI=0.048-1.885 in boys and girls, respectively) and mean academic attainment (β= 0.192; CI=0.021-0.210, β=0.156, CI= 0.049-0.251 in boys and girls, respectively), however, after adjusting by screen time these associations disappeared.
Huang et al. (2015)	Cross-sectional/ Denmark	525 students/12 to 13 age/ 6th to 7th	Andersen intermittent shuttle-run test.	A modified Eriksen flanker task was used to assess inhibitory control. Academic skills were assessed by a math test according to curriculum of Danish schools.	Sex, age, ethnicity, pubertal stage, parental education level, school year, and attending special teaching.	CRF was positively associated with both inhibitory control and math score (β=0.23; CI=0.15-0.32), irrespective of weight status.
Kalantari et al. (2015)	Cross-sectional/ Iran	580 students boys/ 15 to 17 age/ No data grades	BMI, body composition, One-mile walk/run test, sit and reach, hand grip strength test, 40-m sprint, push-ups and Illinois agility test.	Cumulative grade point averages (CGPA) were recorded from the school record of the previous semester (middle of the current year) and end of the current year and divided to 2	PA, pubertal maturation status, age and SES.	CRF (but not PA and muscular fitness) was significantly correlated to better academic achievement in the adolescent boys. Time in the one-mile run/ walk is significantly correlated to CGPA (β=-0.19; p=<0.01).

				for obtaining mean CGPA of the year.		
Sardinha et al. (2016)	Longitudinal/ Portugal	1286 students/ 9 to 14 age (baseline) 12 to 17 age (follow-up)/ 5th to 7th grades.	Progressive Aerobic Cardiovascular Endurance Run (PACER).	Academic achievement was assessed using students' marks at the end of the academic year at baseline and at follow-up 3 year later, in Portuguese, mathematics, foreign language (English), and science.	Age, BMI z-score and AP at baseline.	Students consistently fit increased the likelihood of having high levels of academic achievement in Portuguese (OR= 3.49; CI, 1.97–6.20) and foreign language (OR= 2.41; CI, 1.39–4.14) compared with those consistently unfit. Those that were unfit at baseline and improved their CRF and became fit at follow-up had also higher odds of achieving better marks than those consistently unfit in Portuguese (OR = 2.52; CI, 1.42–4.45) and foreign language (OR =2.13; CI, 1.23–3.67).

CRF= cardiorespiratory fitness; PA= physical activity; BMI= body mass index; AP= academic performance; SES= socioeconomic status; OR= odds ration; CI= confidence interval

Table 2 Risk of bias assessment of the included studies.

	Criteria						
Study	1	2	3	4	5	6	Total Score
Du Toit et al. (2011)	-	-	+	+	-	-	2/6
Kim et al. (2003)	+	+	+	-	-	-	3/6
Eveland-Sayers et al. (2009)	-	-	+	+	-	+	3/6
Wittberg et al. (2009)	+	-	+	+	-	-	3/6
Roberts et al. (2010)	+	-	+	+	-	-	3/6
Davis & Cooper (2011)	+	-	+	+	-	-	3/6
London & Castrechini (2011)	+	-	+	+	-	-	3/6
Van Dusen et al. (2011)	+	-	+	+	-	-	3/6
Wingfield et al. (2011)	+	-	+	+	-	-	3/6
Chen et al. (2013)	+	-	+	+	-	-	3/6
Coe et al. (2013)	+	-	+	+	-	-	3/6
Lambourne et al. (2013)	-	+	+	+	-	-	3/6
Haapala et al. (2014)	-	-	+	+	-	+	3/6
Srikanth et al. (2014)	+	-	+	+	-	-	3/6
Pindus et al. (2016)	+	-	+	+	-	-	3/6
Castelli et al. (2007)	+	-	+	+	-	+	4/6
Chomitz et al. (2009)	+	-	+	+	-	+	4/6
Wittberg et al. (2010)	-	-	+	+	+	+	4/6
Blom et al. (2011)	-	-	+	+	+	+	4/6
Wittberg et al. (2012)	+	-	+	+	-	+	4/6
Bass et al. (2013)	+	-	+	+	-	+	4/6
Kantomaa et al. (2013)	+	-	+	+	-	+	4/6
Liao et al. (2013)	+	-	+	+	-	+	4/6

Rauner et al. (2013)	-	-	+	+	+	+	4/6
Greeff et al. (2014)	+	-	+	+	-	+	4/6
Hansen et al. (2014)	-	+	+	+	-	+	4/6
Scudder et al. (2014)	+	-	+	+	-	+	4/6
Torrijos-Niño et al. 2014	+	-	+	+	-	+	4/6
Chaddock-Heyman et al. (2015)	+	_-	+	+	-	+	4/6
Desai et al. (2015)	+	-	+	+	-	+	4/6
Pellicer-Chenoll et al. (2015)	+	-	+	+	-	+	4/6
García-Hermoso (2016)	+	-	+	+	-	+	4/6
Morita et al. (2016)	-	-	+	+	+	+	4/6
Dwyer et al. (2001)	+	+	+	+	+	-	5/6
Kwak et al. (2009)	+	+	+	+	+	-	5/6
Welk et al. (2010)	+	-	+	+	+	+	5/6
Telford et al. (2012)	+	-	+	+	+	+	5/6
Bezold et al. (2014)	+	-	+	+	+	+	5/6
Esteban-Cornejo et al. (2014)	+	-	+	+	+	+	5/6
Janak et al. (2014)	+	-	+	+	+	+	5/6
Sardinha et al. (2014)	+	+	+	+	-	+	5/6
Aguilar et al. (2015)	+	-	+	+	+	+	5/6
Huang et al. (2015)	+	-	+	+	+	+	5/6
Kalantari et al. (2015)	+	+	+	+	-	+	5/6
Sardinha et al. (2016)	+	+	+	+	-	+	5/6

+indicates that a criterion was satisfied; -indicates that a criterion was not satisfied.

1, Does the study describe the eligibility criteria for participant selection?; 2, Are participants randomly chosen from the population?; 3, Does the study cite the sources and details of the method for measuring academic performance and present the reliability of the instrument for the specific age group?; 4, Does the study cite the sources and details of the assessment methods for measuring physical fitness using reliable methods?; 5, Does the study present the analysis for statistical power and use a statistical method that is adequate to test the hypothesis?; 6, Does the study cite the number of participants for each outcome measurement and does this number represent at least 80% of the total sample?

Table 3 Distribution of the studies that investigated cross-sectional and longitudinal associations between Physical Fitness with academic performance by risk of bias and the level of scientific evidence.

Physical fitness components	Studies that demonstrated association	Study by risk of bias	Association or no association according to study risk of bias	Level of evidence
Cross-sectional				
Cardiorespiratory fitness (n=30)	Yes: 27 (90.0%)	Low: 9 (30.0%) Medium: 20 (66.7%) High: 1 (3.3%)	Low risk of bias: Yes: 7 (77.8%) (positive association) No: 2 (22.2%) Medium risk of bias: Yes: 19 (95.0%) (positive association) No: 1 (5.0%) High risk of bias: Yes: 1 (100.0%) (positive association) No: 0 Total: Yes: 27 (90.0%) No: 3 (30.0%)	Strong evidence of positive association
Muscular strength (n=14)	Yes: 5 (35.7%)	Low: 4 (28.6%) Medium: 9 (64.3%) High: 1 (7.1%)	Low risk of bias: Yes: 1 (25.0%) (positive association) No: 3 (75.0%) Medium risk of bias: Yes: 3 (33.3%) (positive association) No: 6 (66.7%) High risk of bias: Yes: 1 (100.0%) (positive association) No: 0 Total: Yes: 5 (35.7%) No: 9 (64.3%)	Uncertain association
Flexibility (n=8)	Yes: 2 (25.0%)	Low: 1 (12.5%) Medium: 6 (75.0%) High: 1 (12.5%)	Low risk of bias: Yes: 0 (0.0%) No: 1 (100.0%) Medium risk of bias: Yes: 2 (33.3%) (positive association) No: 4 (66.7%) High risk of bias: Yes: 0 (0.0%) No: 1 (100.0%) Total: Yes: 2 (25.0%) No: 6 (75.0%)	No association

Cluster of PF (n=5)	Yes: 4 (80.0%)	Low: 0 (0.0%) Medium: 5 (100.0%) High: 0 (0.0%)	Low risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Medium risk of bias: Yes: 4 (80.0%) (positive association) No: 1 (20.0%) High risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Total: Yes: 4 (80.0%) No: 1 (20.0%)	Positive association
Longitudinal				
CRF (n=7)	Yes: 4 (57.1%)	Low: 2 (28.6%) Medium: 5 (71.4%) High: 0 (0.0%)	Low risk of bias: Yes: 2 (100.0%) (positive association) No: 0 Medium risk of bias: Yes: 2 (40.0%) (positive association) No: 3 (60.0%) High risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Total: Yes: 4 (57.1%) No: 3 (42.9%)	Uncertain association
Muscular strength (n=2)	Yes: 1 (50.0%)	Low: 0 (0.0%) Medium: 2 (100.0%) High: 0 (0.0%)	Low risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Medium risk of bias: Yes: 1 (50.0%) (positive association) No: 1 (50.0%) High risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Total: Yes: 1 (50.0%) No: 1 (50.0%)	Uncertain association
Flexibility (n=2)	Yes: 1 (50.0%)	Low: 0 (0.0%) Medium: 2 (100.0%) High: 0 (0.0%)	Low risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Medium risk of bias: Yes: 1 (50.0%) (positive association) No: 1 (50.0%) High risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Total: Yes: 1 (50.0%) No: 1 (50.0%)	Uncertain association
Cluster of PF (n=3)	Yes: 3 (100.0%)	Low: 1 (33.3%) Medium: 2 (66.7%)	Low risk of bias: Yes: 1 (100.0%) (positive association)	Strong evidence of

	High: 0 (0.0%)	No: 0 (0.0%) Medium risk of bias: Yes: 2 (100.0%) (positive association) No: 0 (0.0%) High risk of bias: Yes: 0 (0.0%) No: 0 (0.0%) Total: Yes: 3 (100.0%) No: 0 (0.0%)	positive association

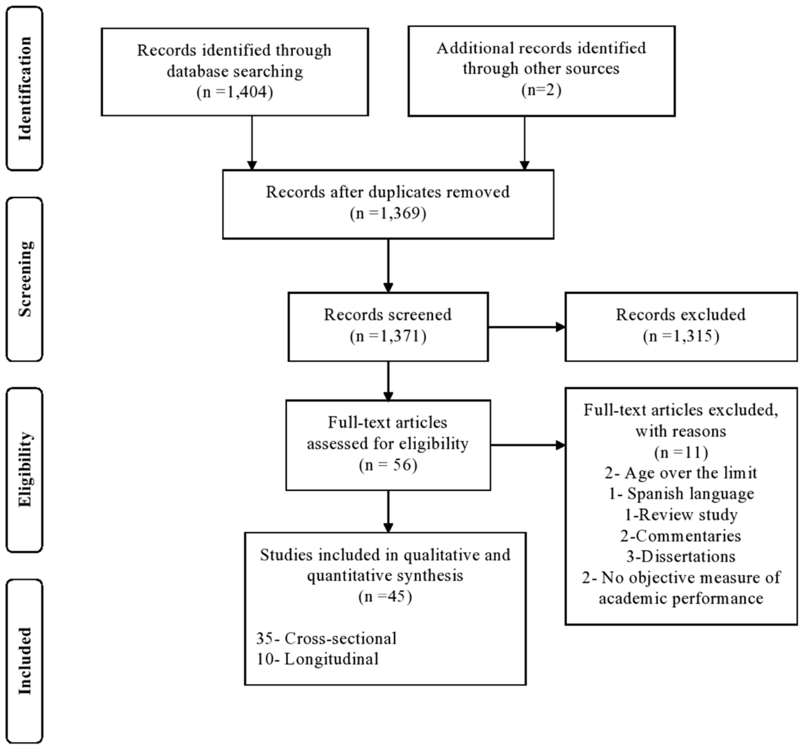

Fig. 1. Flow chart of the study selection process.

449x582mm (72 x 72 DPI)

Appendix 1. Search strategy for MEDLINE (Pubmed)

Independent variable:

"Physical fitness" [DeCS]

"Physical fitness" [MeSH]

"Cardiorespiratory fitness"

"Physical endurance" [MeSH]

"Physical conditioning" [MeSH]

"Muscle strength" [MeSH]

"Muscular endurance"

"Muscular resistance"

Pliability [MeSH]

Pliability [DeCS]

23 ((("physical fitness") OR "cardiorespiratory fitness") OR "physical endurance") OR "physical conditioning"

24 (("muscular strength") OR "muscular endurance") OR "muscular resistance"

25 (flexibility) OR pliability

Outcome:

academic achievement (MeSH)

educational status (MeSH)

attendance school

academic performance

school performance

#26 ((((“academic achievement”) OR “academic performance”) OR “school performance”) OR “attendance school”) OR “educational status”

(#23) AND #26

(#24) AND #26

(#25) AND #26